DOCUMENT RESUME

ED 449 846 JC 010 162

AUTHOR	High, Clennis F.
TITLE	A Broad View of Developmental Program Assessment Activities at Public Junior, Community and Technical Colleges: Evaluating the Evaluators. A Report for Dissemination.
PUB DATE	2000-12-30
NOTE	31p.
PUB TYPE	Opinion Papers (120)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Community Colleges; *Evaluation Methods; Evaluation Utilization; Institutional Characteristics; *Program Effectiveness; Qualitative Research; *Remedial Programs; School Surveys; *Technical Institutes; *Triangulation; Two Year Colleges

ABSTRACT

This study evaluates how various colleges across the country determine the effectiveness of their remedial/developmental programs. The author sectioned the country into 4 quadrants and solicited participation from 80 randomly selected colleges. A total of 50 colleges (63%) responded from over 30 states. Using personal interviews, case studies, document review, and written communications to collect data, the study sought to: (1) add information to the existing literature; (2) examine practicality of a "best practices" with respect to specific college characteristics; and (3) explore evaluations at some colleges that work, but are not often included in the literature. While all colleges considered developmental programs to be important, some colleges (generally those with little or no minority enrollment) did not have developmental programs, while others lacked evaluation procedures for their programs. Four factors were found to impact the quality of evaluations: (1) level of student preparedness when entering college; (2) amount of funds colleges provide to evaluators; (3) knowledge-level of evaluators; and (4) the degree to which some type of accountability factor is associated with the evaluation. The study concludes with 10 recommendations for improving evaluation procedures for developmental programs. (Contains 27 references and an appendix containing respondent college contact information.) (PGS)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 449 846

A Broad View of Developmental Program Assessment Activities at Public Junior, Community and Technical Colleges: Evaluating the Evaluators

A Report for Dissemination

December 30, 2000

Dr. Clennis F. High

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Clennis F. High
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

TC010162

BEST COPY AVAILABLE

Table of Contents

Abstract ----------Page 1

Executive Summary---------- Page 2

Acknowledgements ----------Page 5

Introduction ----------Page 6

Significance----------Page 7

Methodology----------Page 9

Procedures----------Page 10

Objectives----------Page 11

Findings----------Page 12

Recommendations----------Page 20

References----------Page 22

Appendix Section ----------Page 24

Abstract

A study was undertaken to evaluate how various colleges a cross the country determine the effectiveness of their remedial/developmental programs. The focus of the study was not on the outcomes of those evaluations, but how they were done. Drawing from, but not limited to contemporary literature, the researcher sought to explore new avenues of thought and include colleges in areas not normally seen in most current research literature. The researcher sectioned the country into four quadrants and randomly selected fifty colleges from these sections. Care was given to follow-up on non-responding colleges to control "non-response" bias and to be certain the study included a "good mix" of colleges. Using what the researcher calls a "Triangulization" approach he used personal interviews, case studies, document review, and written communications to collect data. The study sought to address three objectives. The objectives included: (1) Add information to the literature not already there, (2) Examine the practicality of the "Best Practice" train of thought when college specific characteristics are considered, and (3) explore evaluation processes at some colleges that work, but is not often included in the literature. After several months of reviewing documents and interviewing college personnel from fifty colleges and over thirty states, the researcher drew conclusions regarding the programs studied, problems at colleges and some state agencies, and identified four areas of concern with the evaluations studied. Moreover, it was found that some colleges do not have evaluation procedures, and many evaluators lack adequate training. Many other important factors are introduced, and ten recommendations are given based on the findings.

Executive Summary

This report was processed, as a means to provide additional information to practitioners who evaluate what is widely known as remedial/developmental programs at junior, community and technical colleges. Moreover, it is intended to take a close look at common beliefs about how programs are or should be evaluated, and the practicality in those beliefs. The report goes a step further and looks at the reality of how programs are actually evaluated throughout the country, and provides a general assessment of the methods and processes used. Lastly, the report addresses the objectives of adding additional information to the literature, examining the practicality of "best practice" thought when college variations and characteristics are considered, and to seek out *fresh* ideas.

The process used to gather data for this report included an extended period of document reviews, case studies, and personal interviews. Evaluation processes were assessed based on acceptable criteria provided by the *Program Evaluation Standards* created by the " Joint Committee on Standards for Educational Evaluations", the *Online Evaluation Resource Library, the Evaluation Primer,* and other notable references. The colleges included in the study from which this report was generated were randomly selected from around the country, based on a systematic process, which greatly decreased the probability of *bias effect*. Of the eighty colleges contacted, fifty or 63% responded and took part in the study. Authoritative sources found in the literature indicate this to be acceptable for this study. Based on what is known about normal distributions, the central limits theorem, and other probability laws these findings are likely to be generalizable. Several state agencies were also included.

The study, which targeted summative evaluations, revealed a variety of activities in regard to how colleges evaluate their remedial/developmental programs. Also it was found that many colleges have no standard method to do this. Because of college specific characteristics, it appears "best practice" methods that works at many colleges are not practical at others.

Four factors seem to impact the quality of evaluations at these colleges. Close examination of the data revealed those to be (1) level of student preparedness when entering college, (2) amount of funds colleges are willing to, or is able to provide evaluators, (3) knowledge-level of personnel conducting the evaluation, and (4) the degree to which some type of accountability factor is associated with the evaluation.

Four broad areas were identified, which outline concerns with the processes used in evaluations, and the most frequent concerns involved under-prepared personnel and methodological errors that threaten the external validity of the evaluation results. The data also showed that many programs are of good quality, and some colleges have high-level researchers. But nearly 20%-25% of those involved needs additional training in research and statistical analysis procedures. Another area of interest was the discovery that many stakeholders are omitted from the planning of evaluations, and dissemination of results often do not reach stakeholders.

Recommendations

Based on careful review of acceptable evaluation procedures, current literature, submitted documents, and personal contacts with individuals at the inclusive colleges and agencies, these recommendations were made.

Recommendation #1: When designing programs at these colleges careful consideration should be given to the make-up of the particular college. Important characteristics in regard to student population, funding ability, personnel, and other significant factors should be paramount in the process. Rather than relying entirely on current literature to design programs, the best should be drawn from the literature and incorporated with institutional characteristics and capabilities.

Recommendation #2: Consortiums of like colleges should be established to exchange ideas and concerns as it relates to evaluating the effectiveness of college remedial programs. The study showed some colleges had very good ideas and practices, while others had different ideas that were also good. Collaboration between these colleges would be beneficial. Common threads exists that would be natural connectors. An example is rural colleges, tribal colleges, urban colleges, and technical colleges. Although there are current consortiums, they mostly are based on geographic locations, and may include colleges with very different characteristics. Contemporary technology will allow interactions of this nature to take place beyond geographic limitations.

Recommendation #3: The greatest amount of concern based on the findings of this study is related to under-prepared personnel. Items A-G under the "inter-process" section show this. All personnel who are charged with conducting evaluations should be unbiased and trained in areas which enhance their ability to conduct reliable and credible studies. It is strongly advised that colleges conduct training workshops using competent college personnel or outside professionals when needed for training potential evaluators.

Recommendation #4: A copy of the *Programs Evaluation Standards, Second Edition: How to Assess Evaluations of Educational Programs* book should be mandatory for for anyone who conducts these evaluations.

Recommendation #5: All stakeholders should be included in the planning of an evaluation. To those who ask who should be included and how much? I would suggest to the degree they are effected by the outcome of the evaluation, it is to that degree they should be included in its design. *The Program Evaluation Standards handbook* gives clear guidelines as to who stakeholders are and how they should be handled.

Recommendation #6: Programs to be evaluated should be well thought out, with considerations given to all impacted by the program. Major changes in the programs should be avoided if at all possible, thereby making the evaluation process meaningful and credible.

Recommendation #7: Colleges should form committees to be involved with and monitor ongoing studies that involve remedial programs, or their independent or dependent measures. This would help decrease duplication of efforts, and provide a continuum process that may establish trend information or possible *control measures.*

Recommendation #8: Colleges should seek to establish some type of control groups for comparison purposes, even if random assignment to those groups is not possible. If there is no control, program effectiveness or ineffectiveness can not be proven with out reservations. This should be easy in Texas where some students in some way delay remediation. If enumerated and grouped together, this can be used as a control group. This was discussed with at least one college. There are other ways to find control measures or expected group standards, books written about experimental research designs are good sources. A book written by Borg and Gall (1989) and the *Online Evaluation Resource Library* are excellent starting places as well.

Recommendation #9: Individuals involved in evaluations should dialog with technical personnel in charge of data collection to increase the quality of data they require to complete the evaluations. Where possible individuals should be periodically assigned to IT departments to develop relationships to help bridge any gaps that exist.

Recommendation #10: Once the evaluation is complete dissemination of finding should be a priority. Students along with other relevant stakeholders should be included.

Acknowledgments

This report that focuses on how community, junior, and technical colleges evaluate their remedial/developmental programs was made possible through the efforts of many people. I have come in contact with many knowledgeable and interesting people across the country. As a researcher; as a person, I have benefited from this experience. From the outset of this study this old researcher new this wouldn't be easy. At the same time I new there was help out there in the way of interested and concerned community, Junior, and technical college personnel. As always, these trusty and knowledgeable individuals came through. This statement is particularly true of counselors, advisers, testers, and other individuals who work or have worked in direct contact with students. **These people are the best-kept secret in the two-year college**. In many cases more information was available from this group than from administrators. Working on the front-line, they are aware of current concerns and trends relating to developmental/remedial education. It seems they are aware of new trends and changes in old ones before it reaches contemporary literature. As one individual in Montana stated: "We live the theory; they write about it".

Some of the more knowledgeable administrators and state agency people have come from a student services background. These individuals were especially helpful in their knowledge as well as their advice. Wayne Stonecypher at the Mississippi State Board for Community and Junior Colleges is such an individual, as is Dr. Estelle Sit, Dean at Laredo College in Texas. Also I would like to thank Dr. Ronnie Davis from Grambling State University for sharing his thoughts about program evaluations.

Knowing that I have probably left someone out, given the enormous amount of contacts I made, I have compiled a list of individuals whom I contacted and interacted with on some level during this process. The procedure required a monumental amount of contacts and document review; consequently, I am certain I missed someone. If that happened, please accept my apology. The list of those individuals can be found in the appendix section. Thanks to you all!

Introduction

Developmental education, as it is sometimes called, seems to have taken on a life of it's own. This statement is even more valid regarding community, junior, and technical colleges. Though the "experts" purport to have *blanket* methodologies and strategies, which if used, will make exemplary programs at virtually any college, this is not necessarily true. There is still much to learn about this area; we can not even decide what it should be called. Individuals at some colleges still call it remedial education, some refer to it as compensatory education, others skill building courses, and of course most now refer to it as "developmental education".

It goes without argument, or it should, that Developmental Education has evolved into a multi-million dollar business. Many colleges in this country are seriously considering "out sourcing" to private companies (Phipps, 1999), which shows there are businesses out there ready to take advantage of this "market". Intire academic departments have been created at universities as a result of this developmental/remedial *phenomenon*, even though the majority of these developmental programs under study exist at community, junior, and technical colleges, not at universities. States, such as Texas and others, bring in experts to tell them things they already know, but for some reason need confirmation. Most states have adequate databases for baseline information of this nature.

A broad base of information is needed to find effective solutions to the remediation issue and assessing its effectiveness. The literature is replete with descriptive statistics, which highlight the major issues, and examples of "best practices", which attempt to address many of these issues. One problem with using a designated "best practice" within the community college arena is the variability found between and within these complex systems. By and large, the overall and *main* purpose of most universities is to accommodate degree-seeking students. Though we have many universities who identify as being *research* institutions, they could not exist without students seeking degrees. This is not the case at typical 2-year colleges.

Today's community, junior, and technical colleges are vastly different from universities, and even between and among themselves (High, 1999). In fact some urban colleges are so widely spread that their inter-college academic cultures, group abilities, program types, and more or very different. Variability exist in the amount of funds received, and the manner in which these funds are distributed among these colleges; the missions and goals of individual colleges, and performance levels of types of colleges. Best practices that work at urban colleges may not be adequate for many rural colleges. Moreover, one would be hard pressed to believe procedures that are favorable for large suburban colleges, with a small number of minorities, would yield the same results at one of the thirty-plus "Tribal Colleges" found in the western and northern parts of this country. Best practice programs are, or should be correlated, with some type of positive/successful outcome. Since initial equivalence is seldom discussed in many of these evaluations, which determine the exemplary programs, the internal validity of these evaluations is often questionable. The results or outcome measures may have been the same with another program or with no program at all. This can be argued by any knowledgeable researcher or statistician, and therefore is a major weakness in

current research and evaluation. If the internal validity is in question, the *external validity*, or degree to which the results can be generalized is likewise questionable. In addition, the evaluation movement is a relatively new one for community colleges. Findings from a study by Hunter Boylan and others (1994) revealed only 14% of 2-year colleges utilized program evaluation in regard to their developmental programs. These are just a few issues that are surfacing and show a need for more information about evaluating college programs; others are given in the next section dealing with the significance of this report.

Significance

For the reading audience two questions quickly come to mind. First, why is a report of this nature important? Secondly, what purpose will it serve? These are basically the questions that many of the individuals contacted asked. Indeed, these are valid questions. While working on a special project in our information technology department, I was required to contact personnel at other colleges in regard to how their programs were set up. When implementing a new system, all ideas are welcome. Through these exchanges, reviewing articles and other literature, and consulting with administrators at my college and others, I realized gaps existed as it related to how developmental programs are evaluated at colleges. Though my institution has a viable program with several measurement methodologies, it was clear this isn't the norm. As I have been involved with similar research in the past, my interest escalated.

Over time this went from an information-gathering endeavor to a research project. At that point I started to nightly send emails to colleges around the country, and called many during evening hours, and during my lunch break. I started to check my email before leaving for work every morning, and sometimes called contacts at 6am (many colleges were in other time zones). When returning home in the evening my first priorities were to check email and phone messages. The responses I got from School Administrators, representatives from state agencies, Presidents, and others were interesting and helpful in seeing this issue in a new and much broader light.

To address question number one, I must turn to logic and statistics. Community, junior, and technical colleges are an important part of our society. Some people believe they are a well-kept secret. Many of the new-millennium jobs require quick and efficient training, and we provide it. Many people want to, at some point, get a degree or certificate; others simply want to add to their skills. Can they do all of this at these colleges? Of course they can. If there were one concept to be used to describe these colleges, I would say it is "ALL PURPOSE". Their campuses are more ethnically diverse than most universities according to an American Association of Community College report (AACC), and provide a myriad of education/training opportunities. Community, junior, and technical colleges hold the trust of the communities they serve, and are preparing the work force (s) of the future. Hence, it follows logical reasoning that well functioning community, junior, and technical colleges are good for the economy, self-realization, families, and society as a whole.

Because, most remediation takes place at these colleges (Yamasaki, 1998, National Center for Education Statistics, 2000), well-designed and efficient programs enhance the overall functioning at these *all-purpose institutions*, and this is desirable. Further it

is not likely developmental education will disappear any time soon. Though there may not be much evidence showing remediation is expanding (phipps, 1998), to my knowledge there is *none* which indicates it is dissipating. In fact, a current report, "The Community College Turns 100", written by the *Policy Information Center* in March states: "Developmental education is a crucial part of the community colleges' commitment to access, student success, and serving the community". That perception is not a foreign concept in the education community; many states share the same view. California for instance recently made an attempt to shift 90% of its remedial coursework to community colleges over several years. While that state had to postpone the measure, Florida was successful in a similar situation, as was New York in the not so distant past. Texas does not have such a policy at this time, but has had many Legislative headaches regarding what should be taught and who should teach it. Moreover, the National Center for Educational Statistics (NCES) in 1995 reported that about forty-one percent of first semester students took remedial reading, writing, and math in community colleges and *ALL* public 2-year institutions offered these classes during that year.

Because many students who are not first-timers also take these classes, the true figure of all students taking these classes is probably considerably higher. Finally, Theodore Martin and Arthur Zeitlin (1993) called remediation a "*phenomenon*" in Higher Education, and stated they expected it to be an important part of the undergraduate curriculum in the future. The prevailing perception seems to be that remedial/developmental education is here to stay, and will largely take place in community, junior, and technical colleges. Adequate evaluation of those programs is and will be essential.

Having made an argument for the importance of these colleges and the continued presence of remedial learning, it is imperative to address something else-money. These colleges are entrusted with billions of dollars of taxpayer money yearly.
It is projected that by the 2006-2007 school year expenditures for public 2-year colleges will reach nearly $27 billion dollars. Because public monies are involved all parties are morally, ethically, professionally, and *probably* legally obligated to do their very best in all areas, including having the best programs possible. It is estimated that remedial education cost public institutions about 1 billion dollars per year (Breneman, and Haarlow, 1998). Though it is argued by some that this amount, compared to the overall cost of education, is relatively small, it is quite a bit of money. In an area already under funded like education, we cannot afford to take 1 billion dollars lightly.

There are many who believe when students take developmental courses, that taxpayers have paid twice for one product. Others (NCES, 1996; Boylan, 1998; Breneman, et al. 1998) disagree, because statistics show most students taking remedial work did not have college-preparatory classes in high school. This may be true as some research revealed this situation takes place in some pockets in public school districts. However, where college remedial programs are not effective, students could pay for classes, or mandatory state tests, twice and even more. Are these students not taxpayers in many of these college districts? It is important that the evaluation process be as "tight" as possible. This is simply doing the best with whatever funds you are given to operate with.

I would vigorously argue that all mentioned above speaks to the importance and significance of endeavors such as this. Continuing that logic I also believe well functioning and productive programs can only help the thousands of students who e r remedial classes each year. Well functioning and productive programs will go far i creating a positive image of these colleges in the public eye. Surely the need for information of this nature is great for these colleges.

Methodology

The information in this report was generated by a systematic procedure. It was a *Triangulization* approach, which combines several avenues of information gathering. The intent was to include a *true* representative group of public community, junior, and technical colleges. The literature suggests there are nearly 1000 of these colleges (American Association of Community Colleges, 1999) located in various types of geographical locations. Because of financial, time, practical, and logistical considerations, only programs at colleges in the continental United States were targeted for review.

Reviewing current literature revealed it often included some colleges, but at the same time often omitted others. For example, there is very little written about remedial/developmental procedures in any of the more than thirty "Tribal Colleges" located in over ten states in this country, or about colleges located in the states of Mississippi, Montana, or South Dakota At the same time colleges, particularly urban and suburban colleges, located in states such as Texas, Florida, California, New York, Arizona, New Jersey, and other large states appear regularly. Many studies and "site visits" have taken place at these locations. It could be that colleges in areas not often reviewed, might hold valuable information that can add to our views on evaluating the effectiveness of remedial programs.

It is the specific intent of this researcher not to systematically omit any type of colleges if at all possible. Table 1, which is located in the appendix section, gives the state agencies, states, and colleges who participated in this study. A list of individuals I contacted is also in that section. While all people contacted were extremely helpful, some people were seriously concerned about the state of evaluation regarding remedial coursework, and exceptionally helpful and supportive in this project.

A most important question surfaced early in the conception of this report. How will data and information be collected? Survey methods and secondary data analyses quickly came to the mind of the researcher, but neither method would provide enough information to determine what processes were being used to evaluate these programs. Indeed the utilization of these methods would follow standard contemporary methodology, but would probably not add much to the pool of information existing in the literature. There had to be some interaction with knowledgeable individuals at these colleges, review of written documents if they existed, an understanding of particular data analysis procedures, as well as some contacts with state agencies and college officials. Moreover, the researcher had to be reasonable and realistic in determining how much document review and interviewing one person could do adequately.

As nearly 95% of these colleges are connected to the Internet (AACC, 1999), this seemed the most cost effective and practical method to approach data collecting. It was

also expected that a great amount of time would be spent on the telephone for purposes of clarification. These communication mediums proved to be a good means of information gathering.

Probably the most fruitful part of process was the phone conversations and face-to-face encounters. During these much information was given and exchanged. One could here changes in the voice of individuals when certain issues were discussed. Although financial constraints made it impossible to travel far, the investigator was able to visit some individuals at surrounding schools. And phone conversations were a constant for agencies and colleges in other areas. This type of contact was so useful in helping to learn what was happening at colleges, it became apparent that it should in someway be part of the recommendation section of this report.

Procedures

The researcher divided the country into quadrants. The matrix's four sections were the Northwest, Southwest, Northeast, and Southeast sections of the country. Twenty colleges were randomly selected for study within each section. The total number of colleges to be contacted was eighty; the total number of colleges responding was 50. The response rate then was 63%. The 63% participation rate is thought to be adequate for this type of research (Borg & Gall, 1989; Earl Babbie, 1978). It should be mentioned here that many studies report 70% rates as being the threshold of acceptability. This mindset is relative to data that will be manipulated through some statistical analysis, such as in survey research. The seventy-percent response rate threshold is desirable and common in quantitative research. This study is based on qualitative methods, comparative analyses, and interviews. We will not attempt to do any data analysis, but do feel the number of randomly selected participants here is adequate to give some since of what is happening at these colleges overall. Further, nine out of the ten state agencies contacted by the researcher responded to the study. This was very encouraging.

Since all colleges did not respond, there are some states not included. However, the researcher reviewed the list of non-respondents to be sure there was not certain types or groups of colleges systematically omitted because of non- response. Where this was found to be the case, extra effort was made to contact colleges in that category a second time. This strategy seems to have worked. An example of this is the "Tribal Colleges." It seemed extremely difficult to get data from this group, but luckily after several follow-up attempts two colleges from that group did contribute *valuable* information. Particular emphasis was also directed toward representation of urban, suburban, and rural colleges. An earlier study by the researcher revealed important differences between them.

Initial contact was made via the Internet asking knowledgeable individuals questions about how they evaluated their remedial/developmental programs, for copies of program procedures, and copies of any institutional evaluations they have completed. Telephone and email was used for follow-up, if it was necessary. The data collection and interview process went on for approximately four months, some information was sent after the planned collection period, consequently it is not included in the analyses or this report.

The final part of the procedure involved a detailed analysis of all data, documents, and personal interviews with individuals. Much time was dedicated to reviewing completed studies and evaluation done at many colleges, and by or for state agencies. For reference, the investigator called upon his personal experience and training, as well as referring to authoritative literature on effective program evaluation and research. Also "The Program Evaluation Standards Handbook" created by The Joint Committee on Standards for Educational Evaluations (JCSEE) was heavily utilized, as was information from the Online Evaluation Resource Library (OERL). Two knowledgeable individuals in the evaluation area were also consulted regarding good evaluation practices. One of these individuals is a statistician who works at Grambling State University where one of the four graduate programs in Developmental Education exist. Online information for colleges regarding evaluation processes was researched. If a process existed online, it was reviewed in dept by this investigator. This process was particularly pain staking, as most of the wording and verbiage in documents that were found were very vague, and spoke to politics and legalities more so than evaluating educational processes. Phone conversations were often required for clarification. Too, many schools submitted copies of their formalized documents, and those documents were carefully read to gain understanding of the process in question. It was also necessary to contact some schools by phone several times, or additional email to be certain nothing was misconstrued, however this was required far less with individual college documents than with state documents.

Summary

After cross sectioning the country into quadrants, colleges were randomly selected from each of the four sections and asked questions about how they evaluated their remedial/developmental programs. Documents obtained from these schools (electronic and hard copies) were reviewed for content, and colleges were contacted to gather additional detail about the programs if needed. Telephone, mail and email were used as forms of contact. In addition, several state agencies were questioned about evaluation practices in their states, and several submitted documents explaining the evaluation process. Because of wording and other concerns, it was necessary to contact some colleges and state agencies to clarify procedures outlined in official documents. Case studies were conducted on some colleges. Finally, all processes were analyzed to determine to what degree they adhered to basic evaluation principles. Several resources were used as references to guide the analysis.

Objectives

The major focus of this study is to increase the depository of information we have on how assessment is done regarding additional learning students are sometimes required to do, sometimes labeled remedial, developmental, or skill-building course work. ***It is important to understand this report focuses on <u>methods used</u> to evaluate programs not the outcomes or results of those programs.*** To that end there are three major objectives connected to this project.

A. To add additional information, which can be used by educators over and above that already known and available in current literature.
B. Examine the practicality of popular beliefs and "best practices" of assessment of remedial/developmental programs across a broad variety of institutions.
C. Ascertain effective processes created by some colleges but not widely publicized in current literature.

Above I have outlined the basic objectives of this study. It is important for the readers of this document and practitioners of remedial evaluation and other educators to understand findings and recommendations are limited to these objectives. So often consumers of research tend to over generalize findings, and researchers are guilty of not communicating obvious limitations of research designs and findings. It is not the intent of this researcher to do that. It is good ethics to acknowledge and be mindful of stated limitations, questions, hypotheses, or objectives in any study.

Findings

After careful review of hundreds of pages of evaluations, reading responses from questions from individuals around the country, talking to individuals by telephone and in person, I have grouped information into two parts; a ***general overview*** and ***specific findings***. The general overview includes a broad summary of information I have collected. I have categorized the ***specific findings*** of this study in to four separate, but interrelated sections. Those are (1) *pre-process factors*, (2) *inter-process factors,(3) post-process factors, and (4) auxiliary factors.* The first group relates to things that were done or should have been done in preparation for the evaluation. Inter-process factors were things that took place during the actual evaluation process or should have been done. These could be methodological matters, measurement and evaluation approach, personnel used and such. The thing here is they take place during the process. Post-process factors are simply things that took place at the completion of the evaluation. These include things such as, reports and reporting; dissemination of findings; explanation of research design, reiteration of evaluation limitations, etc. And finally are what I term as auxiliary factors. Auxiliary factors are the support mechanisms used in these program evaluations. Basically this involved information technology systems used to collect and provide data. These are not necessarily part of an evaluation, but they normally support the evaluation staff by providing data. A general overview of the findings will precede the specific findings.

Overview

Contact with the fifty or so colleges across the country generally showed that community, junior, and technical colleges that have a process, in most cases, are doing a good job in their attempts to evaluate the effectiveness of their programs. All colleges considered evaluating developmental programs as being important, but this occurred in degrees; some colleges and agencies were more enthusiastic than others. For example, Dr. Wayne Stonecypher of the Mississippi State Board of Junior and Community Colleges conducted a quick survey of colleges in his state to see how they approached

evaluating their programs. The colleges were exceedingly helpful in submitting information to Dr. Stonecypher and myself. Many college officials from other states were also eager to reveal their efforts.

Generally, it appears that most research and resulting recommendations come from scholars from the university environment. Though some of these individuals have studied remediation for many years, they are not in the environment on a daily bases. Consequently, many of these colleges are trying to design programs that fit neatly into models devised by these scholars; some are working some are not. The uniqueness, missions, and population characteristics of these two-year colleges are sometimes not considered in these paradigms. This research shows that it is the researcher who is closely associated with the students who flow through these programs that actually see and understand the process the best.

All colleges do not have formal programs, or evaluation processes to evaluate them. Colleges located in areas where there is little or no minority enrollment seemed to have less structured remedial programs, or no programs at all. This was experienced mostly where the minority enrollment was fewer than 10%, and was not uncommon. It must be acknowledged here, however, that this was not necessarily a bad thing. Several colleges in the northwestern areas indicated that most students take college-prep course prior to enrolling. This seems to decrease the need for remediation.

Some colleges like Reading Area College in Pennsylvania use a process that does not necessarily follow standard rules, but seems to work. Faculty members informally discuss the process among themselves on and ongoing basis, and make any adjustments or changes for which they jointly agree. Though this seems more formative than summative, it appears to work for these colleges. Maybe it is because of the level of enthusiasm of the faculty members.

Several states have state-mandated test, those states seemed to have the most structured evaluation programs. These states, like Texas, Florida, and Georgia are monitored on some level regarding outcomes of remedial learning. Because some tests are high-stakes test, the colleges do spend an incredible amount of time developing their evaluation procedures. In interviewing and individual from a large district located in Dade County Florida, it seemed they have their procedures down very well. I reviewed the procedures several times, and questioned the person about standard research procedures normally used in summative evaluations; they were very knowledgeable. Though adhering to most current practices, they constantly evaluate their procedure. Theirs seemed to be very effective. This was not the case for some other large districts.

Colleges with large minority population seemed to be more involved with remediation, but a great number did not have strong methods to evaluate how effective the programs are. It is important to say here that some programs at the colleges were very good, but this seemed to be the exception rather than the rule. Interviewees at some colleges opined that to have good programs, "more funding is needed". Though fully aware of the makings of a good program, some people advised the researcher that funding is a factor that impedes ones ability to create and maintain programs of high caliber. The point was well taken.

Several of the colleges who participated in the study were part of the "tribal college" system. These colleges are located in the north, southwest, and northwest parts of the country. The bulk of the student populations at these colleges are Native-American. Though interviewing individuals from only two of these colleges their message was

almost identical. These colleges (the ones included in this study) relied heavily on knowledge of cultural norms and values, as well as motivating the individual students. In fact, motivation of students in these programs was mentioned several times by the individuals I interviewed. They referred to the classes in the programs as foundation or skill building classes. It seems the small number of students allows for closer contact with them. The individuals interviewed gave many examples of success stories from their colleges. They indicated that the literature ignores or leaves out much of this information about their colleges.

Pre-Process

In review of the documents and interviews with individuals it was determined several problems exist, which occurred prior to the evaluations at many of these schools. One resounding thing was a gap in the level of expertise of evaluators. It seems that many evaluations are done by people with little or no training in the evaluation process, and are limited in their knowledge of applied research methods. Interviews with a number of individuals suggest that they are told to carry out the evaluation and simply do what they are asked to do the best they can. It appears that most feel the evaluations will not go beyond the institution and will not be viewed by the general public. Though well meaning, in many cases this causes a greater problem than doing no evaluation. Periodically knowledgeable individuals outside the "safety" of the institution do view evaluations. The credibility of an evaluation is closely tied to the competence of the investigators. All of the individuals I interviewed appeared to be trustworthy, but that does not make them capable. An example of this can be given. While attending a presentation of an evaluation used to change the test a college used, I witnessed an interesting procedure. The presenter stated that he used scores two-standard deviations above the mean as cut-off points. When asked why this was used rather than one-deviation or a given percentile, or some other method he replied: " I don't know; I'm really not a statistician." Nonetheless, that school elected to change the test used in the program based on this individual's evaluation of that program. Many times student's futures are affected by decisions such as this. Once data is created and presented people quickly forget how and who generated it. Evaluator's credibility is covered in the Program Evaluation Standards, section U2, and is also a potential ethical liability for an institution under certain conditions. About 75% of those interviewed were knowledgeable in regard to the scientific method, evaluation steps and sequences, and how all of this fits together to create a meaningful evaluation.

Where colleges or college systems include multiple sites, there are often important differences between the site populations. Colleges would be wise to keep this in mind when thinking through how their evaluation will be carried out, what will be measured, and how it will be measured. Dr. Carol S. O'Shea, Associate Professor at Owens Community College in Ohio, and president-elect for the National Association for Developmental Education (NADE) believes site-specific populations are an important factor. In giving me her thoughts, she indicated there should be "common threads" in programs, but variations based on relevant population features may well be justified. She also opined that programs should be sensitive to population changes over time, and make changes accordingly.

Too, a large number of evaluators were uncertain of who all the stakeholders were, or so it seems when reviewing their evaluation procedures and documents. Often participating groups such as faculty and state agencies are acknowledged in the evaluat·) process, but not others that may be affected. The most notable stakeholders not mentic l are the students and community. This is particularly important where the program is expected to help student's chances of passing high-stakes test or other obstacles. All reference resources used by this researcher put stakeholder identification and particip)n as a high priority. Section U1 in the Program Evaluation Standards handbook clearly speaks to the importance of identifying and meeting the needs of **all** stakeholders. J: nes Greene (1988) discusses the issue in detail in an *Evaluation Review* article. It would be wise for potential evaluators to read Greene's thoughts; I strongly recommend it.

Four factors seemed to determine the quality of evaluation efforts. Though evaluations ranged from none to very good, these factors were a constant in documents reviewed as well as in conversations and interviews. Those factors include level of student preparedness when entering the colleges, amount of funding available for evaluations, level of expertise of personnel, and whether the evaluation was in some way tied to accountability by state agencies. These factors seemed not only to impact quality of evaluations, but the programs being evaluated too.

Inter-process

Probably the two most noticeable issues during the evaluation process include deviations from acceptable procedures used in ***scientific*** research and haphazardly changing the procedures during the process. The latter of the two was by for the most frequent. This is likely related to the fact that many of these evaluators have had little or no formal training in research and statistics. Given that taxpayer money is used by most of these schools and these programs can impact students' futures, this is a serious matter. I have already acknowledged that many individuals have adequate knowledge to be efficient in carrying out needed program evaluations, but there seems to be a number who don't. In reviewing a huge amount of documents, which included descriptions of data collection, statistical test used, and results of the analyses I have provided a list of common errors found.

A. The sample used is sometimes not representative of the population the results are inferred to. Particularly in large systems that may have distinct sub-populations with totally different characteristics of interest than the system at large.

* *In the appendix section I have included an abstract from a study conducted at a large urban Texas college by Dr. Reynaldo Garay and Mr. Mike Farnell which shows how they approached this.*

B. No indication of why a certain sample size was used. No mention of how the sampling procedure addresses common types of error found when a sample is used in stead of the population, such as sample-error.

C. Not making attempts to isolate the remedial program as an independent variable thereby being able to be relatively sure that **it** was the factor that influenced the identified dependent variable (test scores, retention, persistence, etc.) not some

other intervening or spurious factor. Some schools introduced other factors such as additional group tutoring prior to retakes of high-stakes tests. The question then becomes, was it the program, the additional training session, or an interaction effect between the two which caused the student to do better. It may be that the program was ineffective and the one session was the best method. Or that paring the two provides the best results. There were numerous examples of possible intervening factors in many of the analyses I reviewed.

D. Use of inappropriate statistical tools. The most common practice involved using test scores students made before entering a program and those made after the remedial program to determined if a *significant difference* existed which could be attributed to the program. Here we clearly have "pair or correlated" scores, but many times the analyses revealed that the t-test for "independent" samples was used. Also, several t-test and anova analyses appeared to use data that was arguably not interval, but ordinal level measurements.

E. Little thought given to determine whether a *significant difference* is and *important difference* for the program. Because a statistic is significant does not mean it has educational importance. Significance in a finding simply means you are likely to see the same characteristic in a population. That characteristic may be unrelated to the study are to vague to interpret properly. One study spoke often of an "important significant relationship" between two variables ($r= .23$, $p< .0001$). Though significant, the relationship is weak. In fact, if the author of the study had computed the coefficient of determination (r^2), or the coefficient of non-determination ($1-r^2$) she would have known that the independent variable in this case only accounted for about 5% of the variability in the program dependent variable. What accounts for the other 95%? Other studies compared mean scores that were statistically significantly different, but in reality the huge sample sizes used (i.e. n=1200, n= 900, etc) could be the reason for the difference. This could also be attributed to what is known as the "regression effect". This is a situation where high scores (above average) regress toward the mean, and low scores (below average) move up toward the mean. Unless the "difference" is compared to some other meaningful measure the effect of the program as presented in the evaluation could be misinterpreted. The regression effect is a factor that can bias a study dealing with group comparisons if not addressed properly (Krathwohl, 1988, Tuckman, 1978)

F. The message is often loss in the numbers. About one fifth of the studies reviewed subjected the reader to so many complex statistics and graphics up front, that it is doubtful that all stakeholders *clearly* understood what the evaluation report revealed. In some cases even the researcher could not easily explain it to me.

G. Possible biases in Measures. Most colleges used such dependent variables as, re-test scores, grades in upper-level classes after the program, or retention. re-test scores and retention seem to have "bias potential". Some colleges advised the researcher they created their own tests. Further follow-up revealed that

little thought was given to norming, validity, and reliability of test items or completed test. Tests used on these bases can easily be biased or lead to inconsistent performance. In addition, some schools have use persistence from semester to semester as a measure of retention of program participants. In some cases it appears that people may not return the following semester, because they have completed their goals, not because of anything related to the colleges' programs. Because many schools do not have a process to *clearly* separate student intent, they could very well be unfairly penalizing themselves.

These were most obvious when reviewing material that was submitted and discussing the process with knowledgeable contacts. But also, interviews with individuals revealed that in many cases after programs were set up, **major portions** were changed in the middle of the semester. There is some evidence this may be a result of unclear or unrealistic objectives in the planning stages. Some people indicated they believed this to create problems and confusion in their college's programs. According to current material from the *Evaluation Primer,* many authorities think evaluation programs should have clear, measurable, and realistic objectives to be successful. The greatest concerns voiced by individuals I contacted were these sudden changes were confusing and not fair to some stakeholders, particularly the students.

The problems with analyses are important, because they could yield inaccurate results. Once numbers are crunched and pretty graphics with well-written narratives are produced, most people don't concentrate on how and where the numbers were derived. We, in many cases, fail to realize that research should follow the "scientific method" if the results are to be taken seriously. This fact was reiterated to me in an interview with Dr. Ronnie Davis, a statistician at Grambling State University where one of the few graduate programs in Developmental Education exists. Dr. Davis advised me that when evaluators make claims based on research and statistical procedures, they are held to the same stringent guidelines, as would any researcher. He stressed the fact that training in these areas should be a prerequisite for anyone who is allowed to represent an institution by conducting evaluations, because of the number of stakeholders involved and implications of purported findings.

Post-Process

After the evaluations have been completed it seems the greatest issues involves sharing the results with stakeholders and building on previous research done at the institution or elsewhere. Dissemination of reports that highlight evaluation findings is very important. In some cases, I had to look very hard to find copies of recently completed evaluations and supporting data. In particular, the "full studies" which highlights step by step procedures outlining how the results were obtained were not always available. This should never be the case at any institutions. The process is meant to share with stakeholders and make improvements to programs. It is natural to become suspicious when evaluations are kept "on the down low". Stakeholders are entitled to have access to these reports. The knowledgeable professionals I have spoken with all concur on this issue. Moreover,

Sections P6 and U6 in the Evaluation Handbook clearly address this issue. These sections stress the importance of findings disclosure.

It seems that colleges that had well thought out programs and evaluations were eager to share their work. This was the case even when the evaluations revealed problems that exist in programs. Since this study focused on how evaluations are carried out, not program results, this was a non-issue. It is better for colleges or agencies to share their findings with explanations of why problems occurred with methods to correct them, than to avoid sharing the information.

This is not just a problem with colleges; I ran into similar problems with some state agencies. For example, while reviewing an evaluation and its follow up, supported by the Texas Higher Education Coordinating Board (THECB), I noticed that both studies (two years apart) reported identical response rates of (69%) for community colleges. Since the studies were conducted at different times, I wondered if it was the same 31% group not responding both times? I also questioned if there was something about this group that may be different from the rest of the community colleges in the state, or particular to the group or groups in question? It is common practice to make every effort to obtain a representative sample. With a response rate of 69% a researcher should make an honest effort to get from 5-10% of the non-response group (Tuckman, 1978) to help control sample bias. I was able to get a copy of both evaluations, but no one has responded to the questions mentioned above. I have worked with members of the (THECB) before on other studies, and they have always been very open and helpful. I am still seeking this information.

Secondly here, colleges are not utilizing all of the information they have available. In several cases studies were done prior to a major program evaluation. According to the people I spoke to, credible members of the college community did them. Yet, little or no effort was made to review this information to help better design or enhance the upcoming study. Research is an effort to add to a body of knowledge. It is likely there was important information omitted from the evaluation by not reviewing other work done at the institution pertinent to the program under study.

It is important to understand that the scientific quest for information involves more than collecting data and writing a summary report or executive summary. Depending on what area of study one subscribes to there are five or six steps or chapters. If you make a knowledge claim based on the scientific method, you can expect someone somewhere to ask to see the steps or review the design. If a researcher has clearly identified and followed a particular acceptable process, then he/she has been prudent and responsible. If they have used all resources available at hand, then they are less likely to run into unwanted surprises.

Auxiliary factor

Most colleges seemed to have some type of data collection systems, but some were vastly superior to others. Quality and quantity of data is greatly dependent on the level of technology at colleges. Two colleges, Metropolitan College in Missouri and Lee College in Texas, have recently installed a new system that has excellent capabilities. Individuals I spoke with indicated they are currently making plans to expand their data gathering activities. Other colleges with small budgets, limited personnel, and other limitations are

not so lucky. In some instances, individuals do data collection using lists, and calculators. Though their populations are smaller, many do number well over 1500 and up, and methods like this obviously impact their evaluation procedures.

Another obvious concern is the lack of clear communication between "field people" and "technical people". This seemed the most conspicuous at large urban and suburban colleges. This *digital divide*, if you will, did impact several evaluations. In some cases colleges have an institutional research department as a go between, but they are not always able to be as helpful as they would like it seems; many seem to be over worked. In situations where there was at least one person in the field who understood what data was needed and how to convey that need in technical terms, the programs appeared to fair better.

Lastly concerning auxiliary factors were state agencies. Though not directly involved with evaluations, they should in some way support or guide colleges in what is expected with some clarity, particularly in those states with high-stakes test. Basically the people I interviewed at the 50 colleges and ten state agencies were attempting to do this.

Texas, however, seems to be inconsistent in this endeavor. I spoke to many people and reviewed eight institutional plans for Texas. There has been several changes in this state over the years, and colleges in some cases are frustrated, especially individuals who deal directly with students. The main issues appear to be frequent changes and vague requirements. In reviewing several documents it was revealed that some colleges have institutional requirements that may be greater than those of the state. And the guidelines are so *flexible* that in some situations students who do not pass the state-mandated TASP test are allowed to not take remedial classes for a period but still attend. Because the state has indicated that all incoming students, who are not exempt or waived under certain conditions, must (1) take the TASP test and (2) take remedial coursework in any area not passed it is unclear how program evaluations can be accurate. In addition, it is possible for students to attend "non-course based" remediation such as private tutoring. In the plans I reviewed, many gave no indication as to how these different independent variables could be segregated and then associated with the specified dependent variables (i.e. retention, college course grade, TASP scores, etc.).

Unless all these groups are kept separate in some way, it can not be readily established that it was the program alone that caused variation in students' behavior and academic performance. Moreover, some interviewees stated that some colleges are making changes in program criteria during the semester, and before the next evaluation. They expressed concern that they would not be able to determine if progress is a result of the program before the change, or after the change, and it is causing problems for many stakeholders.

Summary

The findings above are based on careful review of submitted documents, verbal interviews, and/or written communications over several months. The findings are based on content comparisons to acceptable procedures laid out by established authorities in the field of program evaluations, research, and statistical methods. Though there is never a flawless evaluation or study, there are minimal-level criteria that should be present. Care should always be taken to acquire a representative sample, if one is used. All

stakeholders should be considered. Good professional ethics should always be in play. One should always be knowledgeable and competent, or seek the advice of one who is. An evaluation should be conducted as if our professional credibility and our students' well being depended on it, because both do.

Having made a case for programs needing improvement does not mean people are not good at their work. I am fortunate to have been associated with so many capable and knowledgeable individuals over the last few months. The challenges for people working at these community, junior, and technical colleges are many and great, I don't think the educational community or society at large has a full appreciation of this.

The one factor not included in the resources I relied on to complete this study is "enthusiasm". There was no method to measure this in this study, but the evidence of its presence was clear in the responses and support I received from participants. Some of these colleges did not have "best practice" level evaluation designs based on criteria used in current literature, but based on their level of enthusiasm they are approaching that level.

Recommendations

Based on careful review of acceptable evaluation procedures, current literature, submitted documents, and personal contacts with individuals at the inclusive colleges and agencies, the following recommendations are made.

Recommendation #1: When designing programs at these colleges careful consideration should be given to the make-up of the particular college. Important characteristics in regard to student population, funding ability, personnel, and other significant factors should be paramount in the process. Rather than relying entirely on current literature to design programs, the best should be drawn from the literature and incorporated with institutional characteristics and capabilities.

Recommendation #2: Consortiums of like colleges should be established to exchange ideas and concerns as it relates to evaluating the effectiveness of college remedial programs. The study showed some colleges had very good ideas and practices, while others had different ideas that were also good. Collaboration between these colleges would be beneficial. Common threads exists that would be natural connectors. An example is rural colleges, tribal colleges, urban colleges, and technical colleges. Although there are current consortiums, they mostly are based on geographic locations, and may include colleges with very different characteristics. Contemporary technology will allow interactions of this nature to take place beyond geographic limitations.

Recommendation #3: The greatest amount of concern based on the findings of this study is related to under-prepared personnel. Items A-G under the "inter-process" section show this. All personnel who are charged with conducting evaluations should be unbiased and trained in areas which enhance their ability to conduct reliable and credible studies. It is strongly advised that colleges conduct training workshops using competent college personnel or outside professionals when needed for training potential evaluators.

Recommendation #4: A copy of the *Programs Evaluation Standards, Second Edition: How to Assess Evaluations of Educational Programs* book should be mandatory for for anyone who conducts these evaluations.

Recommendation #5: All stakeholders should be included in the planning of an evaluation. To those who ask who should be included and how much? I would sug to the degree they are effected by the outcome of the evaluation, it is to that degree t should be included in its design. *The Program Evaluation Standards handbook* g es clear guidelines as to who stakeholders are and how they should be handled.

Recommendation #6: Programs to be evaluated should be well thought out, with considerations given to all impacted by the program. Major changes in the programs should be avoided if at all possible, thereby making the evaluation process meaningful and credible.

Recommendation #7: Colleges should form committees to be involved with and monitor ongoing studies that involve remedial programs, or their independent or dependent measures. This would help decrease duplication of efforts, and provide a continuum process that may establish trend information or possible *control measures.*

Recommendation #8: Colleges should seek to establish some type of control groups for comparison purposes, even if random assignment to those groups is not possible. If there is no control, program effectiveness or ineffectiveness can not be proven with out reservations. This should be easy in Texas where some students in some way delay remediation. If enumerated and grouped together, this can be used as a control group. This was discussed with at least one college. There are other ways to find control measures or expected group standards, books written about experimental research designs are good sources. A book written by Borg and Gall (1989) and the *Online Evaluation Resource Library* are excellent starting places as well.

Recommendation #9: Individuals involved in evaluations should dialog with technical personnel in charge of data collection to increase the quality of data they require to complete the evaluations. Where possible individuals should be periodically assigned to IT departments to develop relationships to help bridge any gaps that exist.

Recommendation #10: Once the evaluation is complete dissemination of finding should be a priority. Students along with other relevant stakeholders should be included.

References

American Association of Community Colleges (1996). National Community College Snapshot, Washington D.C.: Author.

Borg and Gall (1989). Educational Research: An Introduction, Second Edition. White Plains, NY: Longman.

Babbie, E. (1979). The Practice of Social Research, Second Edition. Belmont, CA.: Wadsworth Publishing Co.

Boylan, Hunter R., Bonham, Barbara S., Bliss, Leonard, B. (1994). Characteristic Components of Developmental Programs. Boone, N.C.: Appalachian State University.

Board of Regents of the University System of Georgia. The Regents Testing Program, (August 2000).

Brawer, Florence B. (1996), Retention-Attrition in the Nineties. Office of Educational Research and Improvement, Washington, DC. (ED393510)

Breneman, D. W. & Haarlow, W. N. (1998). Remedial Education: Cost and Consequences, Washington, D.C. The Fordham Foundation.

Committee on Higher Education and The Committee on Education (1996). An Analysis of Postsecondary Student Preparedness and Remedial Education Needs. Eric No. ED396633.

Educational Testing Service, The American Community College Turns 100 Princeton New Jersey: Author.

Evaluation Primer: An Overview of Education Evaluation (2000). Retrieved November 20, 2000 from World Wide Web: http: www.ed.gov/offices/OUS/eval/primer1.html

Greene, J.G. *(1988)*. Stakeholder Participation and Utilization in Program Evaluation. *Evaluation Review, 12(2), 96-116.*

Haertel, Edward H. (1999). Performance Assessment and Education reform. *Phi Delta Kappan*, 80, p662-668

High, C. F. (1998) An Analysis of the Disparity between Urban, Suburban, and Rural Community, Junior and Technical Colleges: It's Not Just A Minority Problem, Eric No. ED423087.

Hodges, Dawn Z. (1998). Evaluating Placement and Developmental Studies Programs at a Technical Institute: Using ACT's Underprepared Student Follow-up Report. *Community College Review. 26, p57.*

Ignash, J. M. (1997). Who Should Provide Post Secondary Remedial/Developmental Education? *New Directions for Community Colleges. 100, p5-19.*

Joint Committee on Standards for Educational Evaluation (1994). The Program Evaluation Standards: How to Assess Evaluations of Educational Programs, Second Edition. Thousand Oaks, CA: Sage

Koski, W. S. (1998). Replacing Remediation with Acceleration in Higher Education: Preliminary Report on Literature Review and Initial Interviews (Rep. NCPI-4-01) Stanford, CA.: Office of Educational Research and Improvement.

Krathwohl, D (1988). How to Prepare a Research Proposal, Third Edition. New York: Syracuse University Press.

League of Innovation in the Community College (1998). Developmental Education : A Twenty- First Century Social and Economic Imperative. Laguna Hills, CA.: Author.

Lyman, B.C. & Miller, E. (1996). Issues Affecting the Definition of Developmental Education. Retrieved September 25, 2000 from World Wide Web: http://www.umkc.edu/cad/nade/nadedocs/96monpap/plmpap96.htm.

Markus, T. & Zeitlin, A. (1993). Remediation in Higher Education: A "New" Phenomenon? *Community College Review. 13, p13-23.*

McMillan, V. K. (1994). Assessment From the State Perspective. *New Directions in Community Colleges*. 22, p5-11.

National Center for Educational Statistics (1996). Remediation at Higher Education Institutions in Fall 1995. Washington D. C.: Author.

Online Evaluation Resource Library. (2000) Criteria for Sound Evaluation Reports. Retrieved September 25, 2000 from World Wide Web: Wysiwyg://856/http://oerl.sri.com/reports/reportsgloss.html.

Ulschak, F. L., & Weiss, R.G. (1976). The Interpersonal Aspects of Evaluation: A Transactional Analysis Model for Viewing Evaluator-Client Relationships. *Education Technology,* 16(11), 18-35.

Tagg, J. & others (1998). The Decline of the Knowledge Factory: Why Our Colleges Must Change. World & I, June 1998, P. 292-319.

Tuckman, B. W. (1978). Conducting Educational Research, Second Edition. New York: Harcourt Brace Jovanovich, Inc

Appendix Section

In This Section I have included:

1. The names of contacts at the colleges
2. The States and college names
3. An Abstract from a study which shows how two researchers from a large urban system dealt with potential sample bias.

Below are individuals from across the country to which I am indebted for their assistance in the completion of this endeavor. The following page gives participating states, colleges, and agencies.

Abood, Nancy	Johnson, Bill
Allen, Johnny	Johnson, Jane
Ball, James	Lundy, Jack
Bailey, Phil	Luzzo, Darrell
Barney, Ben	Mattox, Ronald
Baxley, Dan	McIntire, Dorothy
Benton, Chris	Morgan, Wella
Beota, Maury	Morris, Catherine
Bevilacqua, Paul	Mpinga, Derek
Boggen, Luther	Nettles, Ronnie
Bullpitt, Dorna	Philpot, Claudia
Burke, Debbie	Perica, Nick
Cade, Becky	Rayborn, Kay
Caffey, Davy	Romer, Brett
Cargill, Jim	Rowe, Gail
Chrestman, Charles	Redone, Richard
Crowley, Judith	Shaw, Jerone
Culbertson, Beth	Sit, Estelle
Clary, Laura	Sparkman, Lavinia
Cook, Lillian	Stonecypher, Sandra
Coniglio, Sharon	Stoncypher, Wayne
Davis, Judith	Tasa, Ken
Derr, Debbie	Topper, Dennis R.
Dreyer, Ron	Thompson, Cheryl
Douglas, Tasha	Valdez, Sonia
Donovan, Sandra	Valiquette, Keith
Emurian, Marrie	Wallace, Roslyn
Gerber, Linda	Walley, Denise
Glasscock, Ronnie	Washam, Shawnalee
Guitierrez, Antonio	Wells, Jayne
Grosz, Karen	White, Rick
Hodde, George	Wright, Ted
Irons, Lynn	

Alabama
Southern Union Community College
Arizona
Dene' College
Arkansas
Rich Mountain Community College
California
College of the Desert
Colorado
Aims Community College
Connecticut
Connecticut Community College-College
Florida
Broward College
Miami Dade Community College
Florida Gulf Coast College
Georgia
Waycross College
Idaho
College of Southern Idaho
Kansas
Kansas City Community College
Maine
Southern Maine Technical College
Massachusetts
Nothern Essex Community College
Maryland
Carroll Community College
Mississippi
Copia-Lincoln Community College
East Central community College
Holmes Community College
Itawamba Community College
Meridian Community College
Miss. Gulf Coast Community College
Northeast Miss. Community College
Southwest Miss. Community College
Missouri
Metropolitan Community College
Montana
Dull Knife Community College
New Mexico
Clovis Community College
San Juan College
Santa Fe Community College
North Carolina
Cleveland Community College
Ohio
Culbertson Community College
Oklahoma
Murray State College
Oregon
Mt. Hood Community College
Pennsylvania
Reading Area College
Rhode Island
Community College of Rhode Island-college
Tennessee
Nashville State Technical College
Texas
Alvin Community College
Austin Community College
Brazosport College
Laredo College
Lee College
Odessa College
Panola College
San Antonio College
Utah
Snow College
Virginia
Southside Virginia Community College
Washington
Clover Park Technical College
Wisconsin
Gateway Technical College
Western Wisconsin Technical College
Wyoming
Western Wyoming Community College
Laramie County Community College

Agencies

California Community College System
City Colleges Of Chicago
Community and Technical College System of Kentucky
Community College of Rhode Island
Connecticut Community College System
Kentucky Community and Technical College System
Minnesota State Colleges and Universities
Mississippi Board for Community and Junior Colleges
South Dakota Board of Regents
Texas Higher Education Coordinating Board

A Tale of Two Colleges:

Abstract

A study was conducted at two colleges in a large urban community college system. The main objectives were (1) to determine which of two measures (objective score vs. writing sample score) was the best predictor of success in the college-level English classes students took at the colleges and (2) to determine what score or interval should be used as a cut-off value for success to take place. In this study, GPA measured success in the classes. The scores were taken from the ASSET test, one of several tests used at these colleges. The test is used as an *alternative* to the state-mandated TASP test, which has been in use in Texas since 1989. Two different, but comparable methods were used at the two colleges to control for the vast diverse population in this system, as well as for other possible intervening or confounding factors. Research has shown that certain college specific characteristics in large geographically spread out colleges can bias studies if not considered and addressed in the design. A large random sample (N = 385) was drawn from the colleges, one hundred seventy two from one and two hundred thirteen from the other. Several inferential and descriptive procedures were conducted on the data, which were clearly conclusive. The analysis for both colleges strongly suggested that the objective scores on the ASSET test were the best predictors of students' success in college English at these colleges when measured by GPA. Also, cut-off scores were established based on analysis of the objective scores.

*** This abstract provided by Dr. Reynaldo Garay and Professor Mike Farnell illustrates the importance to some researchers of obtaining a representative sample when generalization of findings is anticipated. The researchers, knowing that one of these colleges was approximately 85% minority collected a stratified sample that helped adjust for that fact. In speaking with the authors of this study, they indicated that several studies they are aware of failed to account for this "populations within the population" factor, and the findings of those studies later proved to be inaccurate.**